

IR の機能としての学修成果の評価に関する一考察

藤 木 清

I IR について

昨今、大学の中で様々な部局に散在していたデータを1カ所に集約し、大学の経営情報や教育活動のために活用しようとする動きが盛んになってきている。このような活動はIR (Institutional Research) と呼ばれ、また IR に活用するデータはIR データと呼ばれている。

米国で最初に IR が登場したのは、1924年のミネソタ大学といわれており、その後、1960年代に急速に拡大した。1965年には、IR の専門職協会であるAIR (Association for Institutional Research) が設立され、IR 専門職の育成や学会活動が行われてきた。現在、1,500以上の機関が加盟している。

IR の機能は大きく分けて、①政府関係機関への報告 (外部への公表情報)、②財務管理など大学の経営上の意思決定のための情報提供 (経営に関する情報)、③学修成果の評価 (教育に関する情報) の3つがあげられる¹⁾。さらに、それぞれの機能の中でさらに細分化され、多岐にわたっている。

大学の組織体や規模によって様々であるが、上記の役割を遂行するために学長もしくは副学長の直下に評価室やIR 室を設置している例が多い。スタッフは専任教職員を数名擁するところから、兼任の教員と非専任教職員でまかなっている例などがある。

1) (Volkwein 2012)

さて、本稿では、IRの3つの機能のうち、最も重要である学修成果の評価に焦点をあてる。ここで、重要であるという理由は、学修成果の達成状況を評価することにより、内部的には教育改善のデータとして活用することが可能であること、また、外部に対しては、大学全体あるいは学部全体の組織的な教育力をアピールすることが可能であることがあげられる。

II 学修成果について

1. 学修成果とは

日本でIRが注目されるようになってきたのは、平成20年12月24日中央教育審議会「学士課程教育の構築に向けて」の答申が出されたときにほぼ一致する。答申では、各大学が自らの教育理念と目標に基づき、学生の成長を実現する学習の場として学士課程を充実させることを強く求めている。また、それにともない、学生の学修成果の達成状況を評価することにより大学の質保証を充実させ、内外に公表することを求めている。

ここでいう学修成果とは、ある大学あるいは学部学科での学修の結果、学生が理解し、あるいはできるようになることが期待される知識や技術などである。これは、単なる知識の修得にとどまらず、答申の中で学士力として示されているような技術や態度志向性も含まれる。また、大学あるいは学部学科は、学生が達成すべき学修成果を評価可能な表現であらかじめ教育目標として示されることが重要である。

昨今、欧米では、学修成果を重視した教育が進められており、日本の高等教育もその方向に進んでいるのである²⁾。

学修成果を重視した教育を進めるには、個々の教員の教育改善の努力もさることながら、大学全体、あるいは学部学科全体として組織的に教育改善に取り組んでいく必要がある。その試みが学生の学修成果の達成のために有効かどうかを評価するためには、大学全体あるいは学部学科のデータを集約す

2) 川嶋 (2009)

る機能であるIRの活用が重要となってくる。

2. 学修成果の評価方法

学修成果の評価・測定については、様々な模索が続けられている³⁾。しかしながら、どれか1つの評価方法を採用するだけでは不十分であり、複数の評価方法を組み合わせて多角的に評価する必要がある。

学修成果の評価方法は、大きく分けて直接評価と間接評価がある⁴⁾。

2-1 直接評価

直接評価は、学生のテストやレポートの結果を使い、あらかじめ設定した達成目標をどの程度達成しているのかを評価するものである。通常の科目のテストやレポートを使用することは可能であるが、ある特定の学修成果に関する総合的な達成状況を知るためには科目のテストやレポートの採点結果を単純に合計するわけにはいかない。まず、科目のテストやレポートを学修成果ごとの観点に分け、それぞれについて評価を行うための科目ルブリックが必要になる。その上で、複数の科目間で同じ学修成果を統合する作業が必要となる。

科目のテストやレポート以外に、ある学習活動の事前事後にテストを行って変化を見たり、教室外活動や課外活動も含めた総合的な学修成果を評価するために学修ポートフォリオを活用している大学もある。

学修ポートフォリオを用いた学修成果の評価については、米国で多くの大学が実施している。その中でも、カールトン大学とクレムゾン大学は特徴的である。

カールトン大学 (Carleton College) は、ミネソタ州にあるリベラルアーツカレッジであり、様々な Writing Program を実施することにより、学生の文章作成能力を強化している。Writing Portfolio もその一つである。学生は、

3) (笠原2011; 吉田2009) など

4) (洪井他2012)

2 年次において指定された数のレポート等を学修ポートフォリオに保存し、教員がルーブリックを用いて評価を行う。その結果、補習が必要と判断された学生には一定期間、文章作成の補習を受けることが義務付けられているのである。

また、サウスカロライナ州にあるクレムゾン大学 (Clemson University) は、多くの学部を持つ公立大学である。同大学では general education に関する 9 つの学修成果の達成を教育目標として掲げており、そのうちの 8 つの目標については、学生がそれに関連する artifacts (成果物) をポートフォリオに保存し、教員がルーブリックを用いて 5 段階 (0 点～4 点) で評価を行う。8 つの目標うち、7 つは 2 点以上、1 つは 1 点以上 (計 15 点以上) の評価を得ることを卒業の要件としている。同大学では、現在、ポートフォリオ評価とその他の IR データを連結したデータベースを構築し、学生個人の評価と大学全体の評価を行えるよう整備している。

一方、日本においては、学修ポートフォリオを導入している大学は増えているものの、学修成果の評価にまで活用している例はあまり見られない。

そこで本稿では、直接評価の実例として、日本語運用能力に関するテストを用いた評価の一例を後の章で取り上げる。

2-2 間接評価

間接評価は、学生に対する学修行動調査 (アンケート) を利用して、学修の経験や行動、意識などについて定点観測 (例えば、入学時と卒業時) を行い、その変化をもって学修成果を評価するものである。基本的には、統計処理を行い、全体的な評価に用いるが、属性間で比較したり、特定の学習プログラム (インターンシップなど) の参加者の傾向を全体と比較したり、別の IR データと連結して分析に用いることも行われる。そのためには、個人パネルデータとしてデータを集積する必要がある。

米国では、大学間で共通の学生調査を実施して、他大学との比較を通して自大学の特徴を捉え、教育改善に活用することが行われている。事例として

は、米国 IUPUI が主催する NSSE (National Survey of Student Engagement) や、私立大学を対象とした HEDS コンソーシアム (The Higher Education Data Sharing Consortium) が実施している学生調査などがある。日本でも 2012 年度に大学 IR コンソーシアムが設立され、加盟校は共通の学生調査を実施することにより、他大学との比較が可能となっている。

このように、学修成果の評価は様々な方法が開発されている。本稿では、学生調査を用いた間接評価の一例を後の章で取り上げる。

Ⅲ 直接評価の事例—日本語運用能力テスト—

1. 日本語運用能力テストの開発の経緯

文部科学省の戦略的大学連携支援事業の一環として、関西国際大学、神戸親和女子大学、比治山大学から日本語教育関連を専門とする教員が集まり、共通の日本語運用能力テスト（以下、「日本語テスト」という。）を開発した。

関西国際大学では 2009 年度からすでに新生の学習支援を目的に 10 分間で実施できる独自の同テストを開発し、実施していた。

日本語テストの目的は 2 点あげられる。第一の目的は、入学時における日本語運用能力を短時間で測定し、運用能力の優れた学生や学習支援の必要な学生を早期に発見することである。日本語運用能力は大学での学習を行う上での基本的スキルであり、ユニバーサル化を迎えている大学にとって課題の一つとなっている。日本語運用能力は卒業論文やエントリーシートを書くための基本的スキルであり、卒業や進路決定にも深く関連する。そこで、入学時のスコアがローレベルな学生に対しては、ライティングやリーディング等の学習支援プログラムへの誘導が可能となり、ハイレベルな学生には、低学年から高度な教育プログラム等へ誘導することが可能となる。

第二の目的は、定期的に本テストを実施することにより日本語運用能力の伸長を測定することである。昨今、初年次教育が多くの大学で取り入れられるようになっており、その内容は大学で学ぶ意義や自校教育、学習技術など様々である。初年次教育の意義は大学への学習面、対人関係、生活全般への

適応を促すことにあり、特に、学習面への適応として、日本語運用能力を伸長することは初年次教育として重要である。したがって、定期的に日本語テストを実施することにより日本語運用能力の伸長を測定することができ、初年次教育における学修成果の指標の一つとして活用することが期待できる。

戦略的大学連携支援事業の実施にあたり、朝日新聞社が開発したeラーニングシステム「e学び力」を連携校で導入した。このシステムは、新聞記事を素材にした問題等を用いて、聴く力・調べる力・書く力・読む力・整理する力を自学自習で学ぶeラーニングソフトである。また、これらの能力は大学での学びだけでなく社会人としても必要となる基礎的な能力である。

実施の結果、e学び力の参加校全体と比較した場合、連携校3校とも、参加校全体の5つの力の正答率と同じ傾向を示したが、特に、「読む力」「調べる力」が他の力に比較して低い結果となっていた。この結果から、特に、これらの力の基礎となる日本語運用能力について、これを測定するテスト連携校で共同開発し、共通メジャーとしての活用する検討に至ったのである。

2. 日本語運用能力テストの概要

先行して関西国際大学で実施していた「日本語運用能力テスト」は、問題の配付・説明・解説時間が5分、解答時間が10分の計15分で実施していた。連携校で新たにテストを開発するにあたり、この「15分」のメリットを活かし、精選された日本語テストの開発を目指した。この方針の理由は、①被験者が集中して取り組めること、②特別な時間枠や教室準備等の負担なく実施できること、③得点集計と分析とを効率的に行うことができ即応性をもって学生の学習支援に活かせることがあげられる。

パイロットスタディに基づいて検討を重ねてきた結果、最終的な問題の内容は次のとおりである。

- ・問1 漢字の書き取り：5問
- ・問2 漢字の読み：5問

- ・問 3 慣用句（漢字一字を補入）：5 問
- ・問 4 語彙選択（外来語・類義語・待遇に関する語）：5 問
- ・問 5 読解（400字程度の説明文）に関する問題：5 問 計25問（点）

漢字の書き取りと読みの問題では、学生の大学入学までの学習量を測ることを意図している。すなわち、この問題のポイントが高い学生は、大学入学以前から学習習慣が比較的身につけていると考えられ、逆に、この問題のポイントが低い学生の中には、入学以前の学習履歴に何らかの問題を抱えている学生がいると考えられる。

慣用句の空欄補充・語彙選択問題は、語彙量と、表現レベルでの日本語運用能力を測ることを意図している。伝統的な慣用表現は、学生にとっては普段使用することは少なく、語彙量や読書量の推測を意図している。

また、語彙選択問題では、外来語（ビジネス用語や時事用語）は、新聞等において取り上げられている問題にコミットしているか、待遇表現は、同世代を超えた人間関係の中での用語選択が適切に行えるかを問うており、コミュニケーションとしての日本語運用能力を測ることを意図している。

さらに、「連携版」では、文章題を導入することとした。400字前後の平易な説明文を読み、内容を読み取る力、文章を整える力を測れるものへと工夫した。この問題では、読解や問題整理に関する測定を意図している。

以上の内容の検討を踏まえ、パイロットスタディを重ねたうえで、A、B、Cの3パターンの問題の完成に至った。

3. 日本語テストを用いた分析

3-1 日本語テストによる成績の上位層と下位層の弁別

連携校で開発した日本語テストは、2012年度から使用を開始した。2012年度生は入学時にA問題を、2年次9月にB問題を使用した。表1は、関西国際大学の2012年度生全員を対象に実施されたA問題、B問題の結果と1年次10月時点での累積GPAとの相関係数である。入学時のテスト結果と前期終

表 1 日本語テストと GPA の相関 (2012年度生)

	A 問題 (入学時)	GPA (1 年次10月)	B 問題 (2 年次 9 月)
A パターン 入学時	1	.349**	.753**
GPA 1 年次10月	.349**	1	.411**
B パターン 2 年次 9 月	.753**	.411**	1

** 1 % で有意。

了時点の成績の相関係数は0.35であり、ある程度の相関関係が見られる。また、入学時のテスト結果と2年次前期終了時のテスト結果との相関係数は0.75であり、10分間の簡易なテストであるにもかかわらず、相関を維持していることがわかる。

次に、表2は2012年度生の入学時のテスト結果(A問題)と1年前期終了時点の累積 GPA に関する分割表である。

テストが6点以下の学生は GPA 1 未満が25%、GPA 1~2 未満が53%であり GPA の低い学生のうち、約8割の学生を弁別している。一方、テストが19点以上の学生のうち、30%が GPA 3 以上の学生であるため、上位の学生についても、ある程度は弁別できているといえる。

したがって、この結果から、本テストにより成績を完全に予測することは

表 2 2012年度生の A 問題得点 (入学時) と GPA (1 年次10月) の分割表

(単位: %)

		GPA				
		1 未満	1 ~ 2 未満	2 ~ 3 未満	3 以上	総計
日本語 テスト	0 ~ 6	25.0	53.1	12.5	9.4	100.0
	7 ~ 12	14.5	37.7	40.9	6.9	100.0
	13 ~ 18	7.7	30.8	53.8	7.7	100.0
	19 ~ 25	4.4	24.1	40.9	30.7	100.0

できないものの、本テストを用いて成績の上位層と下位層を弁別するという目的はある程度達成できている。すなわち、短時間の簡易テストであっても、学習支援の必要な学生の抽出や日本語のポテンシャルの高い学生の抽出がある程度可能であることがわかる。

4-2 日本語運用能力の伸長

次に、大学入学時と2年次10月時点のテスト結果の伸びを概観する。表3は、A問題とB問題の得点の平均値を、全体、A問題の下位群、A問題の上位群について求めたものである。全体の平均値の差は有意ではあるが、1.1点である。また、問題の難易度による影響もあるので日本語運用能力が一概に伸びているとは結論付けられない。

しかしながら、一般的には上位群に比べて下位群の伸び率が上がる傾向があり、ここでも同様の傾向を示しているため、わずかながら学修成果が表れていることがうかがえる。

表3 A問題とB問題の平均値

	A問題 (入学時)	B問題 (2年次9月)	伸び (B-A)	
全体 n=264	12.5	13.6	1.1	**
上位群 n=132	16.0	16.7	0.7	**
下位群 n=132	8.9	10.6	1.7	**

** 1%で有意。

IV 間接評価の事例－学生調査－

1. 学生調査による評価の概要

次に間接評価の事例について述べる。本章では、本学で実施されている学

生調査のうち、学習面に関連する項目について、1年次11月時点と4年次2月時点の結果を平均値で比較し、学習面における変化について分析する。調査対象者は2009年度生全員である。なお、卒業時の成績にもとづき3群に分けて分析を行っている。なお、当大学では、大学として教育目標として、問題解決力やコミュニケーション力、社会的能動性、多様性理解などが含まれている。

2. 学習面の自信について

表4～表6は、学習面に関する各項目の自信がどの程度あるのかを5件法で回答してもらい、その平均値を比較した結果である。GPA上位群で大きく伸長しているのは、「プレゼンテーション能力」、「人付き合いや対人関係力」、「知性」、「リーダーシップ」であり、コミュニケーション力に関する項目が占めている。

GPA中位群については、「リーダーシップ」が比較的大きく伸長し、「大学の講義の理解度」、「文章作成能力」、「努力をすること」、「教養の深さ」が

表4 自信に関する変化－卒業時累積 GPA 上位群－

	1年次	卒業時	差	
i. プレゼンテーション能力	2.91	3.48	0.58	**
k. 人付き合いや対人関係	3.48	3.98	0.49	**
j. 知性	3.02	3.51	0.49	**
f. リーダーシップ	2.83	3.29	0.46	**
p. 努力をすること	3.68	4.12	0.44	**
l. 自己理解	3.55	3.98	0.43	**
d. コンピュータの活用力	3.22	3.64	0.42	**
m. 文章作成能力	3.19	3.57	0.38	**
g. 数学的思考	2.66	2.96	0.29	**
b. 教養の深さ	3.18	3.44	0.26	**
a. 大学の講義の理解度	3.56	3.78	0.22	*
e. いつも冷静でいること	3.45	3.66	0.21	

**1%で有意、*5%で有意。

表 5 自信に関する変化—卒業時累積 GPA 中位群—

	1 年次	卒業時	差	
f. リーダーシップ	2.84	3.39	0.55	**
a. 大学の講義の理解度	3.14	3.56	0.41	**
m. 文章作成能力	2.85	3.23	0.39	**
p. 努力をすること	3.51	3.88	0.37	**
b. 教養の深さ	3.04	3.41	0.37	**
l. 自己理解	3.50	3.86	0.35	**
j. 知性	2.87	3.22	0.35	**
i. プレゼンテーション能力	2.85	3.20	0.35	**
d. コンピュータの活用力	3.25	3.59	0.34	**
k. 人付き合いや対人関係	3.51	3.77	0.25	*
e. いつも冷静でいること	3.39	3.63	0.24	*
g. 数学的思考	2.61	2.75	0.14	

** 1 %で有意、* 5 %で有意。

表 6 自信に関する変化—卒業時累積 GPA 下位群—

	1 年次	卒業時	差	
i. プレゼンテーション能力	2.52	3.17	0.65	**
m. 文章作成能力	2.77	3.40	0.63	**
l. 自己理解	3.42	3.98	0.55	**
f. リーダーシップ	2.66	3.17	0.51	**
k. 人付き合いや対人関係	3.54	4.05	0.51	**
e. いつも冷静でいること	3.14	3.64	0.49	**
j. 知性	2.76	3.24	0.49	**
d. コンピュータの活用力	3.21	3.65	0.44	**
a. 大学の講義の理解度	3.18	3.53	0.35	**
g. 数学的思考	2.51	2.84	0.34	**
b. 教養の深さ	3.06	3.36	0.30	*
p. 努力をすること	3.51	3.66	0.15	

** 1 %で有意、* 5 %で有意。

続いている。特に、大学の講義や教養など、知識理解に関する項目が上位に入っている点が興味深い。

GPA 下位群では、「プレゼンテーション能力」、「文章作成能力」といったスキル系の項目が大きく伸長し、ついで、「自己理解」、「リーダーシップ」、「人付き合いや対人関係」といった人間関係に関する項目が上位を占めている。

以上の結果から、成績によって項目の違いはあるものの、学習面の自信に関する伸長が確認できる。

3. 学習技術について

次に、学習技術に関する変化について分析する。表7～表9は主に読解や文章作成の学習技術である、文章要約、箇条書き、事実と意見の書き分けについて4件法で回答してもらい、平均値の差をまとめたものである。

表7 学習活動に関する変化—GPA 上位群—

	1 年次	卒業時	差	
1) 本などの内容を決められた字数で要約できる	2.74	3.09	0.34	**
2) 本の内容や話の要点を箇条書きにできる	2.88	3.09	0.20	**
3) 自分の意見と事実を分けて書くことができる	3.02	3.17	0.15	

** 1 %で有意、* 5 %で有意。

表8 学習活動に関する変化—GPA 中位群—

	1 年次	卒業時	差	
1) 本などの内容を決められた字数で要約できる	2.52	2.94	0.42	**
2) 本の内容や話の要点を箇条書きにできる	2.59	2.97	0.38	**
3) 自分の意見と事実を分けて書くことができる	2.81	3.12	0.31	**

** 1 %で有意、* 5 %で有意。

表9 学習活動に関する変化—GPA 下位群—

	1 年次	卒業時	差	
1) 本などの内容を決められた字数で要約できる	2.50	2.86	0.36	**
2) 本の内容や話の要点を箇条書きにできる	2.55	2.82	0.27	*
3) 自分の意見と事実を分けて書くことができる	2.70	2.94	0.24	*

** 1 %で有意、* 5 %で有意。

GPA 上位群は1年次よりスコアが比較的高めであったことから、変化量は少ないが、他の成績群に比べて卒業時のスコアが高めである。

この3つの項目では、「文章要約」の変化が最も大きく、ついで「箇条書き」、「書き分け」と続く。ただし、スコアの高さはこの逆順となっている。

この3つのスキルの中では「文章要約」が最も高度な作業であるため、ほかの2つのスキルに比べてスコアは低めであるが、学習活動を通じて、要約の技術が伸びていると考えられる。

結果として、あくまでも学生の自己評価ではあるが、文章作成や読解に関する能力の伸長が確認できる。

V 今後の課題

IR の視点から見た学修成果の評価の課題として2点あげられる。第一に、評価を教育のPDCA サイクルに活用するために、大学全体あるいは学部全体のアセスメントプランを構築した上で実施する必要があるということである。

アセスメントプランは、各大学においてどのような学修成果を達成するのかについて教育目標として設定することと連動して構築していく必要がある。そのため、教育目標は評価可能な表現になるように具体化することが必要である。

アセスメントプランの策定にあたっては、大学および学部学科、科目、学生個人といった3段階のフェーズを設定する必要がある。それぞれのフェーズで評価材料、評価時期及び評価者を決め、そのプランに従って評価を実施していくのである。最も基本となるのは学生個人のフェーズである。学生一人一人が教育目標として設定された学修成果を達成しているのかを評価するための評価材料を整理する必要がある。次に、科目のフェーズを考える。教育目標を達成するには教育目標の要素が各科目の学習目標として具体化する必要がある。その上で、その科目の履修者がどの程度学習目標を達成しているのかを評価する必要がある。そして、大学及び学部学科のフェーズでは、

個々の学生の学修成果の評価を標本として抽出し、全体としてどの程度達成しているのかを評価するのである。

第二の課題は、学修成果の評価の開発である。これまで述べてきたように、1つの評価方法で正確に測定することは困難である。様々な評価方法を組み合わせることで多角的に評価する必要がある。これまでに多くの評価方法が開発されているが、今後も、可能な限り実態を正確に表現するとともに、実行可能な方法を模索していかなければならない。そのためには、統計的な分析手法を活用することが重要である。

学修成果の評価のみならず、中途退学者のモデル分析など統計的手法を用いた IR データの活用法を開発していくことが重要である。

(筆者は関西国際大学人間科学部教授)

参考文献

- Volkwein, J. F. (2012), "The Structure and Functions of Institutional Research Offices," in Howard, Richard D. et al. ed., *The Handbook of Institutional Research*, John Wiley & Sons.
- 笠原千絵 (2011) 「学修成果の評価方法とルーブリックの活用—アメリカの高等教育団体と大学におけるインタビュー調査から—」『関西国際大学研究紀要』第12号、37-46頁.
- 川嶋太津夫 (2009) 「ラーニング・アウトカムズを重視した大学教育改革の国際的動向と我が国への示唆」『名古屋高等教育研究』8、173-191頁.
- 渋井進他 (2012) 「学習成果に係る標準指標の設定へ向けた検討：国立大学法人評価における評価結果報告書の分析から」『大学評価・学位研究』第13号、3-18頁.
- 吉田文 (2009) 「大学生の学習成果の測定をめぐるアメリカの動向」山田礼子編著『大学教育を科学する：学生の教育評価の国際比較』東信堂.